

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 07-311751

(43)Date of publication of application : 28.11.1995

(51)Int.Cl.

G06F 15/163  
G06F 9/46  
G06F 12/06  
G06F 12/08  
G06F 13/12

(21)Application number : 06-006558

(71)Applicant : BULL HN INF SYST IT SPA

(22)Date of filing : 25.01.1994

(72)Inventor : ZULIAN FERRUCCIO  
RAMOLINI ANGELO  
BAGNOLI CARLO  
LAZZARI ANGELO

(30)Priority

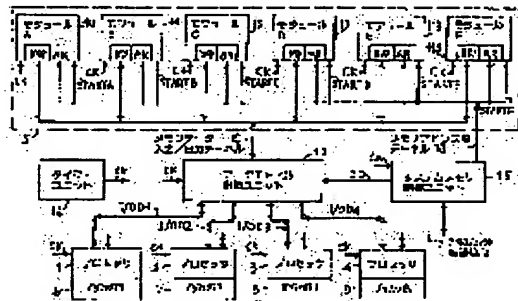
Priority number : 93 93830022 Priority date : 25.01.1993 Priority country : EP

## (54) MULTI-PROCESSOR SYSTEM HAVING SHARED MEMORY

(57)Abstract:

PURPOSE: To execute relatively high speed access to plural processors and a memory, and to guarantee the coherence of data for realizing a high performance data processing system related to a multi-processor system having a shared memory for simultaneously executing plural processings.

CONSTITUTION: Processors 1-4 of plural groups perform access to a module constituting plural shared memories, and the operation of the processors 1-4 and the operation of the module are time regulated by a common synchronizing signal. This system is provided with a system memory control unit 15 in which time is controlled by this synchronizing signal, data channel control unit 16 controlled by the control unit 15, but 17 for address/command transfer, plural data channels I/O (i), channel 18 for a memory address, and memory data input/output channel 19 for operating input/output transfer to the module.



## LEGAL STATUS

[Date of request for examination]

23.01.2001

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or]

BEST AVAILABLE COPY

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平7-311751

(43)公開日 平成7年(1995)11月28日

(51)Int.Cl. <sup>6</sup>	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 F 15/163				
9/46	3 6 0 D	7737-5B		
12/06	5 3 0 F	7608-5B		
12/08	3 1 0 B	7608-5B		

G 0 6 F 15/ 16 3 2 0 M

審査請求 未請求 請求項の数8 O L (全 22 頁) 最終頁に続く

(21)出願番号 特願平6-6558

(22)出願日 平成6年(1994)1月25日

(31)優先権主張番号 9 3 8 3 0 0 2 2 : 5

(32)優先日 1993年1月25日

(33)優先権主張国 イタリア (I T)

(71)出願人 594015026  
ブル アッカエンネ インフォメーション  
システムズ イタリア ソチエタ ペル  
アツィオニ  
イタリア国, トリノ, 10014 カルソ, ビ  
ア マルティリ ディタリア, 3

(72)発明者 フェルッチオ ツリアン  
イタリア国, ミラノ, 20010 コルナレド,  
ピア ジ. ビ. ピコ, 12

(72)発明者 アンジェロ ラモリニ  
イタリア国, バレセ, 21040 チスラゴ,  
ピア ウンベルト イ, 117

(74)代理人 弁理士 石田 敬 (外3名)

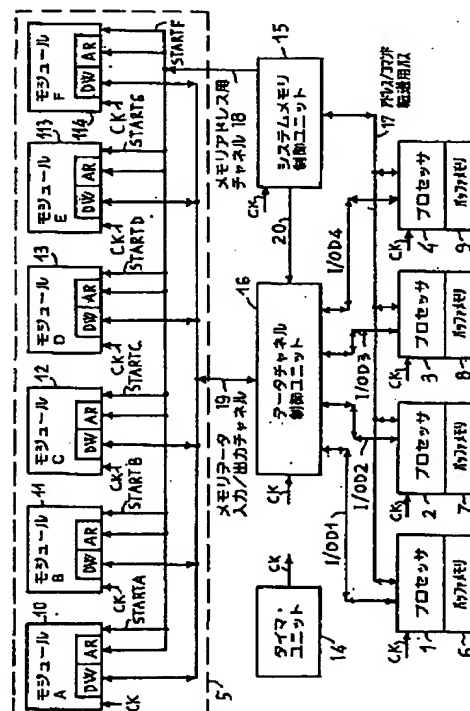
最終頁に続く

(54)【発明の名称】 共用メモリを有するマルチプロセッサ・システム

## (57)【要約】

【目的】 本発明は、複数の処理を同時に遂行させるための共用メモリを有するマルチプロセッサ・システムに関し、高性能のデータ処理システムを実現するために、複数のプロセッサおよびメモリに対し比較的高速のアクセスが実行されると共に、データのコヒーレンスが保証されるようにすることを目的とする。

【構成】 複数のグループのプロセッサ1～4が、複数の共用メモリを構成するモジュールに対しアクセスを行い、プロセッサ1～4およびモジュールの動作が、共通の同期信号により時間規定がなされ、この同期信号により時間が制御されるシステムメモリ制御ユニット15と、制御ユニット15により制御されるデータチャンネル制御ユニット16と、アドレス/コマンド転送用バス17と、複数のデータチャンネルI/O (i) と、メモリアドレス用チャンネル18と、モジュールに対し入力/出力転送を行うメモリデータ入力/出力チャンネル19とを備える。



1

## 【特許請求の範囲】

【請求項1】 各々のグループが少なくとも一つのプロセッサを含む複数のグループのプロセッサ（1、2、3および4）が、複数の共用メモリを構成するモジュール（10、11、12、13、113および114）に対しアクセスを行い、前記のグループのプロセッサ（1、2、3および4）、および、前記の共用メモリを構成するモジュール（10、11、12、13、113および114）の動作が、共通の同期信号である周期的信号（CK）により時間規定がなされるマルチプロセッサ・システムにおいて、

前記周期的信号（CK）により時間が制御されるシステムメモリ制御ユニット（15）と、

前記のグループのプロセッサ（1、2、3および4）および前記システムメモリ制御ユニット（15）に接続され、かつ、前記のグループのプロセッサ（1、2、3および4）と前記モジュール（10、11、12、13、113および114）との間、および、前記のグループのプロセッサ（1、2、3および4）間で転送されるデータを除いて、アドレスおよび動作コマンドを前記システムメモリ制御ユニット（15）に転送する分岐システムバスであるアドレス／コマンド転送用バス（17）と、

相互接続用の複数のロジック回路からなるデータチャネル制御ユニット（16）と、

前記のグループのプロセッサ（1、2、3および4）の各々に対し単一のチャネルを有し、かつ、前記モジュール（10、11、12、13、113および114）のアドレス指定を行うためのアドレスを除いて、前記のグループのプロセッサ（1、2、3および4）と前記モジュール（10、11、12、13、113および114）との間、および、前記のグループのプロセッサ（1、2、3および4）間でデータを転送するための二地点間接続用チャネルである複数のデータチャネル（I/O（i）、例えば、I/OD1、I/OD2、I/OD3およびI/OD4）とを備え、

該複数のデータチャネル（I/O（i））の各々は、前記のグループのプロセッサ（1、2、3および4）の一つを、前記データチャネル制御ユニット（16）に対し個別に接続し、

前記マルチプロセッサ・システムは、さらに、

前記システムメモリ制御ユニット（15）により前記モジュール（10、11、12、13、113および114）のアドレス指定を行うために、前記アドレス／コマンド転送用バス（17）および前記モジュール（10、11、12、13、113および114）に接続されるメモリアドレス用チャネル（18）と、

前記モジュール（10、11、12、13、113および114）を前記データチャネル制御ユニット（16）に連結するために、該モジュール（10、11、12、

2

13、113および114）に対し入力／出力転送を行うメモリデータ入力／出力チャネル（19）とを備え、該データチャネル制御ユニット（16）は、前記複数のデータチャネル（I/O（i））を前記メモリデータ入力／出力チャネル（19）に選択的に接続すると共に、該複数のデータチャネル（I/O（i））自身間を選択的に接続するために、前記システムメモリ制御ユニット（15）により制御され、

前記マルチプロセッサ・システムは、さらに、

前記アドレス／コマンド転送用バス（17）を介して、該アドレス／コマンド転送用バス（17）上の前記プロセッサ（1、2、3および4）に送出されるような順序立ったかつ関連するコマンドおよびアドレスを受信する制御ロジック回路を、前記システムメモリ制御ユニット（15）内に備えており、

該システムメモリ制御ユニット（15）は、

前記コマンドの実行のために必要なリソースと、必要な時間における該リソースの使用可能性とを指定し、

さらに、前記の関連するコマンドおよびアドレスを前記メモリアドレス用チャネル（18）上に転送し、同時に、前記リソースが使用可能な場合に、一つのモジュールを選択するための信号を前記メモリアドレス用チャネル（18）上に転送し、

そして、前記複数のデータチャネル（I/O（i））自身間の選択的な相互接続、および、該複数のデータチャネル（I/O（i））と前記データチャネル制御ユニット（16）における前記メモリデータ入力／出力チャネル（19）との相互接続に関する指示と時間規定を行うことを特徴とするマルチプロセッサ・システム。

【請求項2】 前記データチャネル制御ユニット（16）が、入力データを保持するためのレジスタ（37、42）およびD I 1用レジスタ（49）を有しており、該レジスタ（37）は、該データチャネル制御ユニット（16）と結合する前記メモリデータ入力／出力チャネル（19）に接続されると共に、該レジスタ（42）および該D I 1用レジスタ（49）は、該データチャネル制御ユニット（16）と結合する前記データチャネル（I/O（i））に接続される請求項1記載のマルチプロセッサ・システム。

【請求項3】 前記データチャネル制御ユニット（16）が、該データチャネル制御ユニット（16）から出力されるデータを保持するためのD O 1用レジスタ（44）およびレジスタ（62）を有しており、該D O 1用レジスタ（44）は、該データチャネル制御ユニット（16）と結合する前記データチャネル（I/O（i））に接続されると共に、該レジスタ（62）は、該データチャネル制御ユニット（16）と結合する前記メモリデータ入力／出力チャネル（19）に接続される請求項2記載のマルチプロセッサ・システム。

【請求項4】 前記メモリデータ入力／出力チャネル

(19)が、多数の並列形式の前記データチャンネル(I/O(i))と同じ並列性を有しており、前記データチャンネル制御ユニット(16)が、該データチャンネル制御ユニット(16)と結合する前記データチャンネル(I/O(i))の各々に対し、連続して受信した複数のデータを累算するためのカスケード接続形式の複数のレジスタと、該複数のデータを、前記の出力データを保持するためのレジスタ(62)に転送すると共に、前記メモリデータ入力/出力チャンネル(19)に転送する手段とを有する請求項3記載のマルチプロセッサ・システム。

【請求項5】 前記データチャンネル制御ユニット(16)が、前記メモリデータ入力/出力チャンネル(19)における前記の入力データを保持するためのレジスタ(42)と、前記データチャンネル(I/O(i))の一つに出力されるデータを保持するためのDO1用レジスタ(44)との間で各々結合するマルチプレクサ(31)を有しており、該マルチプレクサ(31)は、前記の入力データを保持するためのレジスタ(42)内に保持されたデータ中の連続する部分を、前記データチャンネル(I/O(i))の一つに出力されるデータを保持するためのDO1用レジスタ(44)に転送する請求項4記載のマルチプロセッサ・システム。

【請求項6】 前記マルチプレクサ(31)が、複数のグループの入力を有しており、該複数のグループの入力の各々は、前記データチャンネル(I/O(i))の一つから入力されるデータを保持するための複数のレジスタの一つと結合する請求項5記載のマルチプロセッサ・システム。

【請求項7】 前記システムメモリ制御ユニット(15)が、前記プロセッサ(1、2、3および4)の前記アドレス/コマンド転送用バス(17)に対するアクセスの調整を行うアービトレーション用ロジック(70)を有する請求項1記載のマルチプロセッサ・システム。

【請求項8】 前記システムメモリ制御ユニット(15)が、前記プロセッサ(1、2、3および4)の複数のグループ中の第1番目の要求に応じて、前記モジュールから読み出されるデータの項目のモディファイを行うための介在要求信号を、前記の複数のグループのプロセッサ(1、2、3および4)から受信し、かつ、前記プロセッサ(1、2、3および4)の複数のグループ中の第2番目から該複数のグループ中の第1番目に供給されるようなモディファイされた一つの単位のデータを転送するように、前記データチャンネル制御ユニット(16)を制御する手段を有する請求項7記載のマルチプロセッサ・システム。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、共用メモリを有するマルチプロセッサ・システムに関する。高性能のデータ処理システムを実現するために、タスクを分割することによって複数のプロセッサが複数の処理を同時に遂行するようなマルチプロセッサのアーキテクチャを利用することは、一般に知られている。

【0002】複数のプロセッサ間での協働を実現するためには、これらのプロセッサ同士が情報やメッセージを交換することが必要であり、かつ、これらのプロセッサが同じデータに対し作用し得ることも必要である。これらのプロセッサは、それゆえに、適切な通信チャンネルによって互いに接続され合うと共に、少なくとも1つの動作メモリに対しそれぞれ接続されなければならない。

【0003】さらに、マルチプロセッサのアーキテクチャの手法は、大きな容量と低いコストを有する動作メモリを提供するが、この動作メモリは、複数のプロセッサの各々の動作時間よりもはるかに長い読み書き用の時間を必要とすることもまた、一般に知られている。このため、プロセッサにより提供される能力を充分に利用することができるように、速度の速い局所的メモリ(Local Memory)、または、ある程度の限られた容量を有するキャッシュメモリが用いられる。このようなメモリの各々は、1つのプロセッサと、個々にかつ独立にアドレス指定可能な複数の動作メモリに接続される。

【0004】このような構成では、アドレス指定可能な動作メモリのスペースは、幾つかのユニット間、または、インタリーピング(Interleaving)の規準によるメモリのバンク間に配分される。このインタリーピングの規準は、幾つかのプロセッサによる複数のメモリへのアクセスにおける「争い(Conflict)」の確率を最小限に抑える。

【0005】比較的高速のアクセスが要求される場合、動作メモリ内に記憶されたデータが繰り返し使用されるように、速度の速い局所的メモリが採用される。しかしながら、この場合は、コヒーレンスの問題が生じてくる。すなわち、英国系(アングロサクソン系)の用語を使用すれば、データの「無矛盾性(Consistency)」の問題が生じてくる。

【0006】また、幾つかのメモリモジュールを採用した場合には、各種のプロセッサと各種のメモリモジュールとの間で相互接続の問題が生じてくる。

【0007】

【従来の技術および発明が解決しようとする課題】従来は、少なくとも部分的に前述の問題を解決するような、下記の2つのアーキテクチャによるアプローチが提示されている。

1) 1つめのアーキテクチャは、「バス」のアーキテクチャ、すなわち、分岐方式による通信チャンネルである。

50 【0008】この場合、システム内のすべてのプロセッ

5

サ、および、すべてのメモリは、単一のシステムバスに接続される。この単一のシステムバスは、タイムシェアリング式のリソースを構成する。このようなリソースに対し、上記のプロセッサ、そして、おそらくは、上記のメモリが、限られたかつ重複しない時間間隔でもって互いに競合し合うことにより、アクセスを行う。

【0009】さらに、システムバスへのアクセスは、各種のユニットの要求に応じ、単一のアービトレーション用ロジック、または、配分されたタイプのアービトレーション用ロジックによって割り当てられる。このようなアクセスの構成では、前もって確立されている規準に従っているため、アクセスにおける争いの問題が解消される。

【0010】この種のアーキテクチャは、基本的に、次のような2つの好都合な点を有する。1番目は、2つのユニットを相互に接続するための動作は、すべてシリアル形式にて行われ、かつ、お互いに決まった順番で実行されることである。このために、通信処理の管理が簡単になる。

【0011】2番目は、システムバスに接続されるすべてのプロセッサが、システムバス上で起こるトランザクション(Transaction)をすべて把握できることである。このために、比較的簡単な「詮索動作(Snooping)」、すなわち、監視機構を用いることにより、リアルタイムにてデータの無矛盾性を保証することが可能になる。

【0012】しかしながら、一方で、上記のアーキテクチャには、次のような限界、すなわち、不都合な点があることを考慮に入れなければならない。すなわち、システムバスの各々のワイアが、多数の入力負荷および出力負荷に接続されている点と、負荷に適した電力を有し、それゆえに、比較的速度の遅いドライバ回路が、種々のワイア上の信号の各々に対して必要になる点である。

【0013】さらに、このような負荷が本質的に有する容量性の性質が、転送され得る信号の周波数を制限する。それゆえに、情報の伝達の速度、すなわち、システムバスの「転送速度」も、負荷の容量性の性質により制限される。幾つかのユニット間の読み書き動作において同じリソースを共用することは、アクセスの争いが増えることを意味し、この結果として、応答の問題が増加する。換言すれば、バスに対するアクセスを待つこと、および、可能性があり、かつ、このアクセスの後に続くような要求された情報の受け取りを待つことが増えてくる。アクセスの応答時間は、メモリユニットの応答の遅さによってばかりでなく、起こり得るアクセスの争いによっても決定される。このアクセスの争いの可能性が高くなればなるほど、バスに沿って重要な情報を転送したりこの情報を持ち出したりするのに必要な時間が長くなる。このために、バスが空いている時間が多くなる。

【0014】1) 2つめのアーキテクチャは、「クロスバ

6

ススイッチ」のアーキテクチャ、すなわち、クロスバー・アーキテクチャによる接続である。この場合、互いに交差する複数の通信チャネルにより、複数のプロセッサおよび複数のメモリが、対になって相互に接続される。そして、スイッチを選択的に閉じることにより、対をなすプロセッサおよびメモリが、選択的にかつ相互に接続される。

【0015】この種のアーキテクチャは、基本的に、次のような2つの好都合な点を有する。1番目は、個々のチャネルにおいて、より多くの対をなすユニットが同時に相互通信を行うことができる。2番目は、マトリクス形式による相互通信により、種々の通信ラインのRC負荷を軽減することができる。

【0016】このような好都合な点により、比較的消費電力の制御回路を用い、比較的高い周波数にてシステムを動作させることが可能になる。この種のアーキテクチャにより達成され得る転送速度は、非常に高い。その理由として、このアーキテクチャにて転送される信号の周波数が比較的大きくなり得ること以外に、多くの同時かつ並列になされる転送が存在することが挙げられる。

さらに、対をなして相互に接続されるユニットは、一般に、幾つかの連続するトランザクションによって保持されており、かつ、このトランザクションのチャネル形成、すなわち、「パイプライン形成」を可能にする。さらに、上記のユニットは、リソースが占有される時間の大部分に対し、応答時間の問題を生じさせることなく達成され得る転送速度をさらに増加させる。

【0017】しかしながら、一方で、上記のアーキテクチャにおいてもなおかつ、次に記載するように、深刻な不都合な点がある。すなわち、多くの対をなす相互接続部における同時転送が、複数のプロセッサ間の「詮索動作」を妨げる点と、幾つかのメモリ内、すなわち、幾つかの記憶ユニット内にデータが複製されるような環境では、データのコヒーレンスの程度が悪くなる点である。

【0018】データのコヒーレンスを保証するために、(少なくともアドレスの)同時転送を否認することが必要である。信号の「経路規定(routing)」や、各構成要素の終結点や、相互接続の管理の問題は、非常に煩雑なものになる。本発明は、上記問題点に鑑みてなされたものであり、高性能のデータ処理システムを実現するために、幾つかのプロセッサおよびメモリに対し比較的高速のアクセスが実行されると共に、データのコヒーレンスが充分保証されるようなマルチプロセッサ・システムを提供することを目的とするものである。

【0019】

【課題を解決するための手段および作用】前記目的を達成するために、本発明の主題を構成するマルチプロセッサ・システムは、複数のグループのプロセッサと、これらのプロセッサと通信する複数の共用メモリを構成するモジュールとを備えている。これらの共用メモリは、個

々にアドレス指定が可能な複数のモジュールにより構成される。モジュールとプロセッサとの通信は、アドレスおよびコマンドを転送するためのシステムバス（すなわち、分岐接続バス）を介して行われると共に、二地点間データ転送用チャンネルを介して行われる。この二地点間データ転送用チャンネルは、各プロセッサをデータ・クロスパー相互接続用ロジックに対し個々に接続する。

【0020】本発明によれば、バスシステムのアーキテクチャの利点と、クロスパーのアーキテクチャの利点とを兼ね備えたハイブリッド方式のアーキテクチャが実現される。このようなハイブリッド方式のアーキテクチャは、同じプロセッサおよびメモリ間の幾つかの転送の際に、順序立ったパイプライン構成を可能にする。

【0021】さらに、上記のハイブリッド方式のアーキテクチャは、個々のプロセッサおよびメモリ間の二地点間データ転送用チャンネルの負荷を軽減する。このために、高い周波数にて動作することが可能になる。さらに、上記のハイブリッド方式のアーキテクチャは、異なるリソースを含むような並列形式の転送を可能にする。

【0022】さらに、上記のハイブリッド方式のアーキテクチャは、メモリへのアクセスが、連続して順番通りに行われることを可能にする。さらに、上記のハイブリッド方式のアーキテクチャにおいて、局所的メモリまたはキャッシュメモリ内でデータが複製される場合、全ての処理過程で、アドレス用チャンネルとデータの無矛盾性に関する「詮索動作」をリアルタイムにて遂行することができる。

【0023】本発明の他の態様によれば、共用メモリを構成するモジュール、すなわち、メモリモジュールが、動作時間の部分的な重ね合わせにより動作するように、これらのメモリモジュールが個々に独立して制御される。それゆえに、これらのメモリモジュールは、独立のメモリユニットとして、システムバスまたはアドレスバスに接続された共通のシステムメモリ制御ユニットを介しアドレス指定がなされる。

【0024】このシステムメモリ制御ユニットはまた、システムバスに対するアクセスを行うためのアービトレーション用ロジックとして機能する。このようにして、複数のプロセッサおよびシステムメモリ制御ユニットに対するアドレスバスの負荷は軽減される。本発明のさらに他の態様によれば、データ・クロスパーのロジック、すなわち、データチャンネル制御ユニットは、共用メモリおよびプロセッサの両方に対し、入力／出力レジスタを備えている。

【0025】カスケード形式で一つのレジスタから他のレジスタへ転送を行う構成では、幾つかの転送は、並列に行うことができる。さらに、データ・クロスパーが、単一のデータチャンネルを介してのメモリとのデータ交換を行うための収集部として機能する場合であっても、転送時間の部分的な重ね合わせにより、メモリに対する

「パイプライン形成」が可能である。

【0026】このようなチャンネルは、データの転送速度を制限することのないノードを形成する。なぜならば、ノードを介してのデータ転送に必要な時間は、転送速度の制限内に収まる程度に充分短いからである。本発明のさらに他の態様によれば、相互接続用ロジックは、バッファ用レジスタ（または、バッファ）以外に、メモリおよび各種のプロセッサに対する接続に応じて異なる並列性を有するようなチャンネルを備えている。さらに詳しくいえば、メモリとデータ・クロスパーとの間では、 $N \times M$ バイトであるのに対し、データ・クロスパーとプロセッサとの間では、たった $N$ バイトである。

【0027】すなわち、メモリとデータ・クロスパーとの間の情報転送は、 $N \times M$ バイトのブロックに対し同時行われる。これに対し、データ・クロスパーとプロセッサとの間の情報転送は、各々の期間で $N$ バイトのデータブロック中の1ブロックを転送させることにより、 $M$ 個の連続する位相にてシリアル形式で動作を続けることで実行される。

【0028】このようなシリアル形式の転送は、応答の問題を生じさせない。なぜならば、データ・クロスパーとプロセッサとの間の接続は、一方向性のものであり、相互干渉が起こらないからである。上記の構成では、プロセッサの並列性に比べてメモリの並列性が相対的に高いので、より高速で動作させるためのあり得るプロセッサの要求に対し、メモリ容量の一部またはそのすべてをあてがうことができる。これと共に、各種の電気的構成部品またはユニットの端子の数、および、種々のユニット間の受動接続を、許容され得る上限内に収めることができる。

【0029】このような端子の数の制限は、経済上の都合、すなわち、多数の入力／出力端子を有する電気的構成部品の工業的な実用性によってのみでなく、標準の通信バスの使用が可能なインタフェースを有するような製品として使用可能な電気部品を用いることの便利さによっても、付与されるべきものである。実際に、インタフェースのレベルにおいて、本発明のマルチプロセッサ・システムの主題の基礎をなすような前述のハイブリッド方式のアーキテクチャでは、例えば「VMEまたはFUTURE BUS」タイプの一般的な標準バスが使用されている。

【0030】

【実施例】本発明の特徴および利点は、下記に示すような添付図面を参照しながらの発明の好適実施例の説明から、さらに明確になるであろうと思われる。以下、添付図面（図1～図7）を用いて本発明の実施例を詳細に説明する。図1は、本発明の一実施例に従って構成されるアーキテクチャおよび共用メモリを有するマルチプロセッサ・システムを示す概略的なブロック図である。

【0031】図1のシステムは、複数のプロセッサ1、



2、3および4を備える。これらのプロセッサ1、2、3および4には、それぞれ、バッファメモリ6、7、8および9が設けられている。さらに、図1のシステムは、複数のモジュール10、11、12、13、113および114（おそらくは、モジュールの数は、プロセッサの数よりも多いであろう）により構成されるシステム・メモリ5と、予め定められた周波数のタイミング信号を生成するタイマ・ユニット（TIM UNITと略記されることもある）14とを備える。なお、図1では、上記モジュール10、11、12、13、113および114を、それぞれ、モジュールA、モジュールB、モジュールC、モジュールD、モジュールE、およびモジュールFと表示している。

【0032】さらにまた、図1のシステムは、共用メモリならびにシステムバスのアービトレーションを制御するためのシステムメモリ制御ユニット（SMCユニットと略記されることもある）15と、ロジック回路からなるデータチャネル制御ユニット16、すなわち、データ・クロスバー（DCBと略記されることもある）とを備える。

【0033】プロセッサ1、2、3および4は、一緒にして接続され、さらに、アドレスおよびコマンドを転送するためのアドレス／コマンド転送用バス（ACBUSと略記されることもある）17を介してシステムメモリ制御ユニット15に接続される。このアドレス／コマンド転送用バス17の適切なワイヤを介し、かつ、一般のアービトレーションおよび通信プロトコルを用い、上記の各プロセッサは、バスに対するアクセス要求信号ABREQ（図3）をSMCユニット15に送る。さらに、この各プロセッサは、バス許可信号ABGRANT（図3）を個々に受け取る。その後、このバス許可信号ABGRANTは、アドレス／コマンド転送用バス17を有効に占有し、さらに、メモリアドレスと下記のような複数の信号をSMCユニット15に送る。これらの複数の信号とは、例えば、読み出し、書き込み、または、別の種類の動作（例えば、図3のRWIM）のような、要求される動作を識別するための信号である。

【0034】システムバスであるアドレス／コマンド転送用バス（ACBUS）17は、分岐方式の通信チャネルを構成する。ただし、必ずしもそうである必要はないが、おそらくは、バスに対するアクセス要求信号ABREQ、対応するバス許可信号（バス許可応答）ABGRANT、および、種々のプロセッサの状態信号は、例外になるであろう。この場合、プロセッサの状態信号は、好ましくは、プロセッサの各々とユニット15との間で、二地点間接続方式により交換がなされる。

【0035】ユニット15は、メモリアドレス用チャネル（MADDRと略記されることもある）18を介して、読み書き用アドレスと、この後に続く適切なタイミング・コマンド（STARTA、STARTB、STA

RTC、STARTD、STARTE、およびSTARTE）をシステム・メモリ5に転送する。このタイミング・コマンドは、アドレスに応じて、各種のモジュール（メモリモジュール）10、11、12、13、113および114中の一つを選択し、始動させる。

【0036】これらのモジュール10、11、12、13、113および114の各々においては、アドレスがチャネル（MADDR）18上に存在する時間が、ある程度制限されている場合でも、レジスタARが、必要な時間のすべてにわたって読み書き用アドレスを保持する。また一方で、データの転送は、二地点間接続により行われる。この二地点間接続は、プロセッサ1、2、3および4の各々と、メモリデータ入力／出力チャネル（MDATと略記されることもある）19との間で、あるいは、対をなすプロセッサ間で、データチャネル制御ユニット（DCB）16により、ユニット15から受信されるタイミング・コマンドに基づき選択的に形成される。

【0037】さらに、モジュール10、11、12、13、113および114の各々では、レジスタDWが、書き込むべきデータの1単位を保持する。このようなデータは、書き込み動作に必要な時間のすべてにわたってメモリデータ入力／出力チャネル（MDATと略記されることもある）19から受信される。図1においては、プロセッサ1、2、3および4は、それぞれ、複数のデータチャネルI／OD1、I／OD2、I／OD3およびI／OD4を介してデータチャネル制御ユニット（DCB）16に接続される。

【0038】システム全体の動作は、同期形式で遂行される。この場合、各種のユニットは、すべて、タイマ・ユニット14により生成される周期的信号CKに基づきクロック制御がなされる。図2は、図1のアーキテクチャのデータチャネル制御ユニットの具体的な構成例を示す概略的なブロック図である。ここでは、図1のデータチャネル制御ユニット16を集積回路により構成している。なお、これ以降、前述した構成要素と同様のものについては、同一の参照番号を付して表すこととする。

【0039】ここで、データチャネルが一つの集積回路として形成される程度にこのデータチャネルの類似性が充分高い場合は、データチャネル制御ユニット16は、同じ構成の複数の集積回路として形成され得る。これらの複数の集積回路は、一般に知られている「ビットスライス構成」の概念、すなわち、ビット群によるロジック回路の分割に従って作製される。

【0040】データチャネル制御ユニット16は、基本的に、下記の5種の構成要素を備える。1つめの構成要素は、データチャネルI／OD1、I／OD2、I／OD3およびI／OD4よりデータをそれぞれ入力するための4つのグループの受信部21、22、23および24である。

【0041】2つめの構成要素は、データチャンネルI/OD1、I/OD2、I/OD3およびI/OD4上にデータを取り込むための4つの制御回路、すなわち、ドライバ25、26、27および28である。3つめの構成要素は、メモリデータ入力/出力チャンネル19上にデータを取り込むための単一のグループのドライバ29である。

【0042】4つめの構成要素は、メモリデータ入力/出力チャンネル19からやって来るデータをデータチャンネル制御ユニット16に入力するための単一のグループの受信部35である。5つめの構成要素は、5個のマルチプレクサ30、31、32、33および34である。

【0043】マルチプレクサ30の入力は、4つのグループの受信部21、22、23および24の出力に接続される。さらに、上記マルチプレクサ30の出力は、単一のグループのドライバ29に接続される。このような接続を行うことにより、ドライバ29がイネーブルの状態になった場合に、メモリデータ入力/出力チャンネル(MDAT)19に対し複数のデータチャンネルI/OD(i)の中の一つを選択的に接続することが可能になる。ここで、I/OD(i)中の記号(i)は、便宜上付加しているだけであり、省略されることもある。あるいは、既述のように、I/OD(i)の代わりにOD1、I/OD2、I/OD3およびI/OD4のように表すこともある。

【0044】その他のマルチプレクサ31、32、33および34の各々は、データチャンネルI/OD(i)の中の一つと関係し、かつ、4組の入力を有する。さらに、これらの入力の各々は、受信部35、21、22、23および24の出力に接続される。ただし、この場合、各受信部がそれぞれ関係するデータチャンネルI/OD(i)を有する受信部の出力への接続は除外される。

【0045】さらに、上記マルチプレクサ31、32、33および34の出力は、それぞれ、ドライバ25、26、27および28の入力に接続される。このような接続を行うことにより、メモリデータ入力/出力チャンネル(MDAT)19をデータチャンネルI/OD(i)の中の一つに接続し、かつ/または、おそらくは同時に、2つのデータチャンネルI/ODと一緒に接続することが可能になる。

【0046】マルチプレクサおよびドライバの動作は、デコーダ36により生成される適切なコマンドSEL1、…SELNに従って制御される。この場合、これらのコマンドに対し、周期的信号CKに基づきクロック制御がなされる。ここで、例えば、次のようなことが可能となる点に直ちに注意すべきである。すなわち、データの衝突なしで、データのソースとしてのデータチャンネルI/OD1が、メモリデータ入力/出力チャンネル(MDAT)19と、他のデータチャンネルI/OD(i)の中の一つに接続されるか、または、データのソースとして

のデータチャンネルI/OD1が、2つのデータチャンネルI/ODと一緒に接続されるかし、さらに、第3のデータチャンネルI/ODがメモリデータ入力/出力チャンネル19に接続されることである。図3は、図1のアーキテクチャのシステムメモリ制御ユニットの具体的構成例を示す概略的なブロック図である。ここでは、システムメモリ制御ユニット15に接続されるシステムバスのアービトレーションの構成も一緒に例示することとする。この場合も、システムメモリ制御ユニット15を集積回路により構成することができる。

【0047】図3において、システムメモリ制御ユニット15は、システムバスに対するアクセスを調整するためのアービトレーション用ロジック(ABUS ARB UNITと略記されることもある)70と、有限状態ロジック72(STATEMACHINEと略記されることもある)と、一対のレジスタ73、74と、デコーダ75と、論理和(OR)回路76とを備える。

【0048】通常のタイプのアービトレーション用ロジック70は、その入力において、種々のプロセッサ間の二地点間接続方式により、バスに対するアクセス要求信号ABREQ(i)(記号(i)は、通常、省略される)を受け取る。さらに、ごく一般的な方法を用いて、周期的信号CKにより制御されるタイミングに従い、複数の二地点間接続の一つに応答のバス許可信号ABGRANT(i)(記号(i)は、通常、省略される)を送り込むことにより、システムバスへのアクセスを許可する。このバス許可信号ABGRANT(i)の送り込みは、種々のプロセッサに対し、一連のタイムベースにおける一つの期間毎に行われる。

【0049】上記アービトレーション用ロジック70は、好ましくは、システムメモリ制御ユニット15の集積回路の一部であるが、公知の方法に従ってプロセッサ全体に配分されるアービトレーション用ロジックに置き換えることもできる。この場合、アービトレーション用信号は、分岐接続方式により交換することができる。上記ユニット15は、システムバスであるアドレス/コマンド転送用バス(ACBUS)17を介し、遂行すべき動作を規定するコマンド信号を受け取る。特に、このコマンド信号として、要求される動作が読み出し動作であるか、または書き込み動作であるかを示す信号RWと、読み出されるデータの単位をモディファイ(Modify)するという意図の下での読み出し動作を示す信号RWIMが挙げられる。実際に存在するような他のコマンドは、本発明の発明の範囲外にあるので、それらをすべて理解する必要はない。

【0050】これらのコマンドがシステムバスに転送された後に、どこで動作が遂行されるべきかを示すメモリアドレスが転送される。ここで、プロセッサがバスへのアクセスを獲得した後においてのみ、コマンドおよびアドレスがシステムバスに送り込まれることに注意すべき



である。さらに、他の動作の遂行に既に関係しているかもしれないようなリソース（例えば、メモリモジュール）を、共同して使用することができることに注意すべきである。

【0051】この場合、リソースが空き状態になるのを待つ間にシステムバスが占有されたままになるのを避けるために、システムメモリ制御ユニット15は、コマンドおよびアドレスの内容を分析した後にリトライ信号RETRYに応答する。このコマンドが拒否された場合、要求中のプロセッサは、上記のコマンドを再提示するように案内される。

【0052】このようにして、上記のコマンドは、必要なリソースが使用可能となるときのみ実行される。このために、コマンドが実行される場合には、関係するリソースの実行速度に依存するような予め定められた時間で実行されることが保証される。したがって、メモリからデータを読み出す場合に、このメモリから供給されるデータの順番は、コマンドが受け入れられた順番と同じ順番になる。

【0053】システムメモリ制御ユニット15から受信されたコマンドおよびアドレスは、レジスタ73に保持される。このレジスタ73は、周期的信号CKに基づきクロック制御がなされ、かつ、デコーダ75により復号化される（デコーダ75の入力は、レジスタ73の出力に接続される）。基本的に、上記デコーダは、アドレスおよびコマンドに基づき、どのモジュール（モジュールA、モジュールB、モジュールC、モジュールD、モジュールE、またはモジュールF）を使用すべきか、および、要求されている動作が書き込み動作（書き込み信号R）であるか否かを決定する。上記デコーダはまた、アドレスに応じて、メモリに対し予め定められていないデータの転送を指定する。ただし、信号I/Oにより指定されている複数のプロセッサの一つは、例外とする。

【0054】デコーダからの出力信号は、有限状態ロジック72に伝達される。この有限状態ロジック72は、周期的信号CKに基づきクロック制御がなされる。さらに、上記の有限状態ロジック72は、周期的信号CKの各周期に対し、前に受信した信号の関数として進行する。既に述べたように、リトライのメカニズムの結果としてプロセッサにより要求される各動作が実行される場合には、この各動作は、予め定められた時間で実行される。それゆえに、有限状態ロジック72は、ある時期に受信した信号に基づいて動作することが可能になり、この結果として、現在のクロック周期とこれに続くクロック周期におけるリソースの状態の痕跡を保持することができる。

【0055】それゆえに、有限状態ロジック72は、その出力において、イネーブル信号ENを提供する。このイネーブル信号ENは、予め定められた必要な時間の期間で必要なリソースが使用可能になる場合にのみ、レジ

スタ73に存在するアドレスおよびコマンドを出力側のレジスタ74にローディングすることを可能にするものである。

【0056】レジスタ74は、アドレスおよびコマンド以外に、信号A、B、C、D、EおよびFによってもローディングがなされる。ある時期においては、信号A、B、C、D、EおよびFの中の一つのみが権利を主張する。そして、この一つの信号がメモリアドレス用チャンネル(MADDR)18上のシステム・メモリ5に送られたときに、この一つの信号は、相互に排他的な方式により、複数のモジュール中の一つを選択して始動させる（始動信号STARTA、STARTB、STARTC、STARTD、STARTE、およびSTARTF）。

【0057】さらに、始動信号により開始したメモリの動作に応じて、有限状態ロジック72は、チャンネル20を介し、データチャンネル制御ユニット16（図1）を制御するために適切に時間調整がなされたコマンドを転送する。読み出し動作の場合には、最終的に、コマンド（排除信号）OENA、B、C、D、EおよびFに応じて、選択されたモジュールが、メモリデータ入力/出力チャンネル19上に読み出し後のデータを転送することが可能になる。

【0058】このような結果は、読み出し動作の期間において、「詮索動作」の後に「介入(Intervention)」が起こらないという条件下で生じ得る。このことは、これから考察することとする。キャッシュメモリによるデータ複製機能を有するマルチプロセッサ・システムにおいては、データの無矛盾性は、基本的に、下記の2つのアプローチにより保証される。

【0059】(1) 1番目のアプローチ…モディファイがなされた各データを、即刻メモリ内に書き込むこと、すなわち、ライトスルー(Write Through)。

(2) 2番目のアプローチ…機会が生じたときのみ、モディファイがなされた各データを延期形式にて書き込むこと(ライトバック(Write Back)、またはコピーバック(Copy Back))。

【0060】1番目のアプローチは、データの1単位がプロセッサ内のキャッシュメモリによりモディファイされる度に、メモリへの書き込みを要求する。すなわち、この1番目のアプローチは、バスおよびメモリ・リソース（例えば、メモリモジュール）を相当な期間使用することを意味する。したがって、このようなアプローチは、実用上好ましくない。

【0061】2番目のアプローチにおいては、すべてのプロセッサが、メモリに送られてくる読み出しの要求を監視することにより、モディファイされた形でキャッシュメモリ内に存在するデータの1単位が上記の読み出しに関係するか否かをチェックすることが前提条件となる。この場合、モディファイされた結果として更新されたデー

タのコピーは、キャッシュメモリ内に存在しない。

【0062】上記の2番目のアプローチでは、モディファイされたデータが存在するキャッシュメモリを有するプロセッサは、現在の状況を他のプロセッサに通知し、さらに、要求されたデータを必要とするプロセッサにデータを送らなければならない。そして、このプロセッサのデータは、対応するメモリ内で、送られてきたデータに置き換えられる。このときに、コマンド(排除信号) OENA、B、CおよびDを送出しないことで、メモリの出力が阻止される。

【0063】好ましくは、システムメモリ制御ユニット15は、2番目のアプローチにより動作するようになっている(しかしながら、このユニット15は、1番目のアプローチにより動作するように容易に調整できる)。この2番目のアプローチでは、プロセッサ間の「詮索動作」の信号の交換が簡単に行える。このような動作を遂行するために、システムメモリ制御ユニット15は、種々のプロセッサから、二地点間接続を通して、状態信号 SNOOP OUT (i) を受け取る。これらの状態信号 SNOOP OUT (i) は、種々のプロセッサから、適切なタイミングにより送られる。上記の状態信号 SNOOP OUT (i) は、システムバス(ACBUS)上に存在する読み出し要求が、キャッシュメモリ内にないデータに関係しているか(SNOOP OUT=NULL(データなし))、または、キャッシュメモリ内に存在して有効であり、それゆえに、少なくとも一つのメモリと共用するようなデータに関係しているか(SNOOP OUT=SHARED(共有))、または、キャッシュメモリ内に存在し、かつ、メモリ内に含まれるデータに関してモディファイされるようなデータに関係している(SNOOP OUT=MODIFY(モディファイ))ことを通知することを目的とする。

【0064】上記の状態信号 SNOOP OUT (i) はまた、次のような理由により、「詮索動作」を遂行することが不可能になることを示すこともできる。例えば、プロセッサが動作中であるという理由か、または、プロセッサ間でデータが転送される場合は、転送データが受信できないという理由が考えられる。上記のいずれの理由によっても、トランザクションが完了しないために、このトランザクションを繰り返すことが必要である(SNOOP OUT=RETRY(リトライ))。

【0065】これらの信号は、有限状態ロジック72により受信される。この有限状態ロジック72は、システムの状態と、制御の対象となる動作とを規定する際に、これらの受信した信号を考慮する。これから詳細に述べることはあるが、受信した信号が「MODIFY」を示す場合には、プロセッサは、モディファイ動作を遂行する必要があることをシステムメモリ制御ユニット15に確認した後に、データチャネルI/OD(i)上にデータの1単位を提供するために介在しなければなら

い。

【0066】さらに、有限状態ロジック72は、チャネル20を介し、データ・クロスバー(DCB)の種々の点間で確立すべき接続を適切に制御する。既に進行中のトランザクションにおいてリソースを使用する際に争いが生じた場合には、上記の有限状態ロジック72は、介在の要求に対し最も高い優先権を与える。上記の有限状態ロジック72は、リトライ信号RETRYを提示することにより、現在のトランザクションを停止させ、さらに、動作が繰り返されなければならないことを通知する。

【0067】さらに、種々のプロセッサから受信された状態信号 SNOOP OUT (i) は、論理和回路76内で一緒にされる。この論理和回路76はまた、必要であることが提示された場合には、有限状態ロジック72からリトライ信号RETRYを受信する。さらに、上記の論理和回路76は、出力信号ARESPを生成する。この出力信号ARESPは、システムバスの分岐接続を通して種々のプロセッサに転送されると共に、NULL、SHARED、MODIFY、またはRETRYに対応するようなシステムの可能な状態を表示する。

【0068】図4は、図1のマルチプロセッサ・システムの動作を説明するためのタイミング図である。ここでは、図1のマルチプロセッサ・システムの動作、特に、図2のマルチプロセッサ・システム内のデータチャネル制御ユニット16の動作に関するタイミング・ダイアグラムを簡潔な形で示すこととする。さらに詳しくいえば、図1中の周期的信号CKのダイアグラムは、時間の推移に対する周期的信号(すなわち、クロック信号)CKの状態およびレベルを表している。

【0069】アクセス要求信号ABREQ(i)のダイアグラムは、種々のプロセッサがシステムメモリ制御ユニット15に送ることができるようなアクセス要求の状態を表している。このダイアグラムは、プロセッサの一つに関し、幾つかの通信ラインの電氣的なレベルを表示するという意味において、累積的なものである。

【0070】同様に、バス許可信号ABGRANT(i)のダイアグラムは、システムメモリ制御ユニット15によって種々のプロセッサに送られる応答信号の状態が、時間に対しどのように変化するかを累積的に表している。アドレス/コマンド転送用バス(ACBUS)のダイアグラムは、アドレス、および、このアドレスに関連するコマンド(読み出し/書き込み)を規定する信号の状態の変化を表している。これらのアドレスおよびコマンドは、プロセッサの各々から、互いに異なる時間の期間でシステムバス上に転送される。

【0071】状態信号 SNOOP OUT (i) のダイアグラムは、種々のプロセッサからシステムメモリ制御ユニット15に送られる信号の状態に関し、時間に対する累積的な変化を表している。この変化は、アドレス/

コマンド転送用バス (ACBUS) 上に存在するアドレスに対し続けられる継続的な監視の結果として見い出される。

【0072】出力信号ARESPのダイアグラムは、システムメモリ制御ユニット15からアドレス/コマンド転送用バス (ACBUS) の2つのライン上に送出される信号の状態に関し、時間に対する累積的な変化を表している。これらの信号は、アドレスおよび状態信号SNOOP OUT (i) の受信にตอบสนองして生成される。上記の信号に基づき、システムメモリ制御ユニット15は、動作の実行に必要なリソースが、要求中の時間の期間では使用不可能であり、それゆえに、再度トランザクションを要求することが必要になるために、現在関係するトランザクションが完了しない旨をすべてのプロセッサに知らせるようにしている。あるいは、上記のシステムメモリ制御ユニット15は、現在のトランザクションが、幾つかのプロセッサにより共用されないデータに関係しているか (NULL)、または、共用されるデータに関係しているか (SHARED) か、または、プロセッサによりモディファイされるデータに関係しているか (MODIFY) をすべてのプロセッサに知らせるようにしている。

【0073】さらに、システムメモリ制御ユニット15は、予め定められた優先順位の規準に従い、かつ、トランザクションの実行に必要なリソースの時間的な使用可能性に応じて、複数のプロセッサに対する単一のアクセスを許可する (例えば、かなり前ではあるが一番最後にアクセスを獲得しているプロセッサに対し)。メモリアドレス用チャネル (MADDR) のダイアグラムは、システムメモリ制御ユニット15をシステム・メモリ5に接続するためのメモリアドレス用チャネル18の状態を表している。

【0074】最後に、データチャネルI/OD (i) のダイアグラムは、種々のデータチャネルおよびデータチャネル制御ユニット16の状態に関し、時間に対する累積的な変化を表している。認識され得ることではあるが、周期的信号CKは、複数の連続する時間期間、すなわち、クロック周期P1、P2、…P13を規定する。このクロック周期においては、周期的信号であるクロック信号は、最初にレベル“0”であるか、または、確定したレベルになっている場合 (ロジックのレベルと電気的なレベルとの間の関係は、一切ない) には、レベル“1”に変化する。

【0075】図4においては、クロック周期よりも大きくない時間期間内で種々の信号が提示されたり、消失したりする。さらに、各周期の真ん中におけるクロック信号のレベル“0”からレベル“1”への遷移は、信号の状態が安定になってストロブ、すなわち、信号の認識が可能になる瞬間を意味する。上記の約束事に基づき、システムのユニット間で可能な種々のトランザクション

が、どのように進行するかを検査することが可能になる。

【0076】これらのトランザクションには、基本的に、下記の4つのタイプがある。

(1) データの項目を読み出すために、あるプロセッサi (iは正の整数) により行われるシステム・メモリ5へのアクセス： このタイプのトランザクションは、プロセッサによりアクセス要求信号ABREQ (i) を提示し、その後、アドレスおよび読み出しのコマンドを送ることにより始動する。

【0077】(2) データの項目を書き込むために、あるプロセッサiにより行われるシステム・メモリ5へのアクセス： このタイプのトランザクションは、アクセス要求信号ABREQ (i) を提示し、その後、アドレスおよび書き込みのコマンドを送ることにより始動する。

(3) 他のプロセッサY (Yは正の整数) により始動する読み出しのトランザクションにおいて、あるプロセッサiにより行われる介在： この介在は、メモリから読み出されるデータを置き換える際に、プロセッサYに対しデータの項目を供給する目的で遂行される。

【0078】さらに詳しく説明すると、このタイプのトランザクションは、データの項目がモディファイされ、かつ、プロセッサi内で使用可能である旨を、状態信号SNOOP OUT (i) のラインを介してシステムメモリ制御ユニット15に通知することにより始動する。

(4) I/Oメッセージ、すなわち、プロセッサ間で直接行われる通信： このトランザクションにおいては、あるプロセッサI (Iは正の整数) が、例えば、周辺機器に対する制御機能を遂行するような他のプロセッサYへ直接にデータの項目を送る。

【0079】このタイプのトランザクションは、アドレスによってメモリの外側のスペースが指定されると共に、プロセッサ (または、信号I/O) が特定されるという理由のみにより、書き込み動作と異なる。ここで、一つの例として、周期P1において一つの (または二つ以上の) アクセス要求信号ABREQ (i) が提示されているような図4のダイアグラムを詳細に考察することとする。

【0080】アービトレーション用のシステムメモリ制御ユニット15がアクセス要求を受け取った場合、このユニット15は、周期P2においてバス許可信号ABGRANT (1) を提示することにより、プロセッサ1に対するアクセスを許可する (図4中のABGRANT (i) の最初の#1)。このアクセスは、予め定められた優先順位の規準に従い許可される。例えば、かなり前にバスへのアクセスを獲得したプロセッサに対しアクセスが許可される。

【0081】プロセッサ1がバス許可信号ABGRANT (1) を受信した場合、このプロセッサ1は、例えば

モジュールAを指定するためのメモリアドレスをアドレス/コマンド転送用バス(ACBUS)上へ送出する(周期P3)。システムメモリ制御ユニット15は、このアドレスを受け取り、モジュールAが空いていることを確かめる。すなわち、モジュールAが、既に、読み出し動作や書き込み動作に関与していないことを確かめる。そして、受け取ったアドレスをメモリアドレス用チャネル(MADDR)18上へ送出することにより、モジュールAを始動させる。このモジュールAの始動は、適切なモジュール始動信号およびモジュール選択信号を生成することによって実行される。

【0082】さらに、一つの例として、周期P4の期間でアドレス指定がなされることにより始動するモジュールAは、その後の周期P7の期間で、読み出された情報をメモリデータ入力/出力チャネル(MDAT)19上へ出力する。換言すれば、上記の例において、読み出しサイクルは、その動作を実行するために4つのクロック周期を必要とする。

【0083】周期P7の期間では、システムメモリ制御ユニット15は、モジュールAからの出力を可能にする。さらに、システムメモリ制御ユニット15は、チャネル20を介してのデータチャネル制御ユニット(DCB)16からの出力を可能にする。この場合、メモリデータ入力/出力チャネル19をデータチャネルI/OD1に接続することにより、モジュールAの出力側からプロセッサ1にデータが転送される。このようにして、プロセッサ1により要求される読み出し動作が完了する。

【0084】周期P4からP7までの期間では、モジュールAに対する他の読み出し動作または書き込み動作が遂行され得ないことは、明らかである。さらに、周期P4の期間では、他のモジュールのアドレスを指定する目的でメモリアドレス用チャネル(MADDR)18を使用することは不可能である。同じように、周期P7の期間では、メモリと他のデータチャネルI/ODとの間で他のデータを転送する目的で、メモリデータ入力/出力チャネル(MDAT)19およびデータチャネル制御ユニット(DCB)16を使用することも不可能である。

【0085】上記のように占有されたりソースの状態は、システムメモリ制御ユニット15の有限状態ロジックにより考慮される。しかしながら、一度モジュールA内で読み出し動作が開始されると、メモリアドレス用チャネル(MADDR)18は、空き状態になる。このために、プロセッサ1とモジュール7との間で開始されるトランザクションの確認を完了させるために何が適しているかを考慮する前に、モジュールB、モジュールC、モジュールD、モジュールE、またはモジュールFに係する他の動作が始動可能になる。

【0086】周期P3の期間では、システムバスであるアドレス/コマンド転送用バス(ACBUS)17上に存在するアドレスが、システムメモリ制御ユニット15

によってのみでなく、プロセッサ2、3および4によっても受信される。これらのプロセッサ2、3および4は、同じアドレスにより指定される情報が、それぞれのキャッシュメモリ内に存在するか否か、そして、この情報がどのような形(共有、モディファイ等)で存在するかを検査するために配置される。

【0087】もし、このような情報が存在しないか、または、共有されているのみであるならば、種々のプロセッサは、周期P4の期間において、システムメモリ制御ユニット15に対し、対応する表示(データなし/共有:NULL/S)を有する状態信号SNOOP OUT(i)を送り込む。さらに、周期P5の期間において、システムメモリ制御ユニット15は、NULL/Sの表示がなされた出力信号ARESPをすべてのプロセッサに送ることにより、種々のプロセッサにおいてキャッシュメモリの状態に対する更新の動作が要求されることを確認する。

【0088】ここで、周期P3の期間において、読み出し動作に関し、プロセッサ2は、システムバスへのアクセスを許可されるものと仮定する。周期P4の期間において、プロセッサ2は、モジュールA(2>A)の読み出し動作のために、システムバス上にアドレスを送出する。このアドレスは、モジュールAの読み出しサイクルをたった今始動させたばかりのシステムメモリ制御ユニット15により受信される。

【0089】このような構成によれば、システムメモリ制御ユニット15が、モジュールAにより構成されるリソースが使用可能ではないことを検査した場合に、上記のユニット15は、メモリアドレス用チャネル(MADDR)18上にアドレスを転送しない。さらに、上記のシステムメモリ制御ユニット15が、プロセッサ2、3および4から、読み出し動作の際に、キャッシュメモリ内に含まれるデータの項目が入っていない旨の確認を受け取った場合に、上記のユニット15は、読み出し動作が実行されず、プロセッサ2が読み出し要求を繰り返し提示しなければならない旨をすべてのプロセッサに通知する。

【0090】それゆえに、周期P7の期間において、プロセッサ2は、アクセス要求信号ABREQ(2)を再提示し、さらに、周期P8の期間において、システムメモリ制御ユニット15は、バス許可信号ABGRANT(2)を再提示する(この場合、より高い優先順位をもつ要求が、他のプロセッサにより同時になされることはない)と仮定している)。

【0091】さらに、周期P9の期間において、プロセッサ2は、アドレス/コマンド転送用バス(ACBUS)上にアドレスを再度送り込み、モジュールAに対し読み出し動作を要求する。この場合、必要とされるリソースが空いているので、次のような動作が遂行される。

【0092】システムメモリ制御ユニット15からメモ

リアドレス用チャネル (MADDR) 18上へアドレスが転送される (周期P10)。そして、周期P13の期間において、プロセッサ2から、メモリデータ入力/出力チャネル (MDAT) 19、データチャネル制御ユニット (DCB) 16およびデータチャネルI/OD2を介して、要求したデータの項目が受信される。

【0093】ここで、プロセッサ3がバスへのアクセスを獲得した場合に、このプロセッサ3は、周期P5の期間において、読み出し動作のためにモジュールCに向けられたアドレスをアドレス/コマンド転送用バス (ACBUS) 上に送出する。この場合、モジュールCが空いているので、システムメモリ制御ユニット15により読み出し動作が開始され得る。そして、この読み出し動作は、既に説明がなされている時間的な流れに従って行われる。この時間的な流れは、必ずしも繰り返す必要はない。この理由として、モジュールCから読み出されるデータの項目は、プロセッサのキャッシュメモリのいずれにも存在しないという仮定がなされていることが挙げられる。

【0094】また一方で、あるキャッシュメモリ内にデータの項目が存在し、かつ、モディファイされている場合、トランザクションは、次のような異なる形で進行する。例えば、周期P6の期間において、プロセッサ4がアドレス/コマンド転送用バス (ACBUS) へのアクセスを獲得したと仮定した場合に、このプロセッサ4は、モジュールBに向けられたアドレスをアドレス/コマンド転送用バスに送出する。

【0095】システムメモリ制御ユニット15は、メモリアドレス用チャネル (MADDR) 18上にアドレスを転送し (周期P7)、モジュールBを始動させる。さらに、上記のシステムメモリ制御ユニット15は、状態信号SNOOP OUT (i) に基づき、要求しているデータの項目が、他のプロセッサのキャッシュメモリ内に存在する旨の表示を受け取る (例えば、状態信号SNOOP OUT (3) によれば、プロセッサ3は、モディファイ (MODIFY) の状態にある)。

【0096】それゆえに、システムメモリ制御ユニット15は、「ARESP=MODIFY」の表示 (周期P8) により、アドレス指定がなされたデータの項目が、メモリから供給されずにプロセッサから供給される旨をすべてのプロセッサに通知する。プロセッサ3は、すべての要求が承認されたことを認識する。さらに、周期P10の期間において、システムメモリ制御ユニット15は、プロセッサ3から、データチャネルI/OD3、データチャネル制御ユニット (DCB) 16およびデータチャネルI/OD4を介してプロセッサ4へ、モディファイされたデータを転送することを可能にするような形でデータチャネル制御ユニット (DCB) 16を制御する。また一方で、信号出力の提示を排除するためのコマンドOENBの作用により、モジュールBから読み出さ

れたデータの項目は、モジュールの出力側からは転送されない。

【0097】好ましくは、プロセッサ3からの出力データはまた、既に存在するデータの項目を置き換える目的でモジュール内に書き込むために、モジュールBにも転送される。考えられ得る最後のタイプのトランザクションは、書き込みのトランザクションである。

【0098】例えば、周期P8の期間において、プロセッサ1は、システムバスへのアクセスのためにアクセス要求信号ABREQ (i) を提示する。ここで、より高い優先順位を有する他のアクセス要求が全くない場合には、プロセッサ1は、アドレス/コマンド転送用バス (ACBUS) へのアクセスを獲得する (周期P9、バス許可信号ABGRANT (1) が提示される)。

【0099】それゆえに、周期P10の期間において、プロセッサ1は、モジュールBに対しアドレスを送出し (1>B)、かつ、アドレス/コマンド転送用バス (ACBUS) 上に書き込みコマンドを送出する。システムメモリ制御ユニット15がリソースに関する争いを確認しないという仮定の下に、周期P11の期間において、上記アドレスが、メモリアドレス用チャネル (MADDR) 18に転送される。さらに、書き込むべきデータの項目が、データチャネルI/OD1からメモリデータ入力/出力チャネル (MDAT) へ転送される。

【0100】ここで、リソースが使用可能でない場合か、または、メモリモジュールが動作中であるという理由がある場合か、または、周期P11の期間でメモリデータ入力/出力チャネル (MDAT) 19が動作することが予想されるという理由がある場合 (モディファイ信号MODIFYの後に動作するであろう)、データの項目およびアドレスの転送は阻止されるであろう。さらに、周期P11の期間において、システムメモリ制御ユニット15は、リトライ信号RETRYを提示するであろう。

【0101】この場合、他のプロセッサにより提示されたモディファイ要求 (MODIFY) と一緒に提示される書き込み要求は、次のような2つの異なる方法で処理することが可能なことは、いうまでもないことである。まず第1に、各々のモディファイが提示された場合に、対応するデータの項目が、書き込み動作以前にメモリ内で更新されるように決められているときは、書き込み要求は、モディファイ要求と衝突する。しかしながら、この場合、モディファイ要求が書き込み要求よりも高い優先順位を有しているために、この書き込み要求は、システムバスへのアクセスを許可されない。

【0102】また一方で、たった今、モディファイ信号MODIFYによる他のプロセッサの介在を生じさせたばかりの読み出し動作が、モディファイを意図した読み出し動作RWIT (すなわち、読み出しデータが今後モディファイされるであろうことは、既に知られている)



である場合、メモリ内でのデータの項目の更新は、無意味なものとなる。それゆえに、一つのプロセッサから他のプロセッサへデータの項目を転送するために、複数のデータチャンネルの一つに対するアクセス要求が同時に差し出された場合でも、書き込み要求に関しシステムバスへのアクセスを許可することが可能になる。

【0103】換言すれば、2種の要求の衝突による争いが生ずることなく、2つのデータの時間的な重ね合わせが実現される。このようなことは、一般のシステムのバスアーキテクチャでは、不可能であろう。この時間的な重ね合わせはまた、プロセッサ間のデータ交換動作と、メモリの読み出し動作との間で、リトライ信号RETRYを用いてアクセスの衝突を解消することにより可能である。

【0104】上記の仮定によれば、例えば、周期P10の期間において、プロセッサ3が、データ書き込みのためのアクセス要求を提示し（アクセス要求信号ABREQ(3)が提示される）、かつ、周期P11の期間において、システムメモリ制御ユニット15が、システムバスへのアクセスを許可した場合（バス許可信号ABGRANT(3)が提示される）、プロセッサ3は、プロセッサ1に対し向けられた信号I/Oと同じ信号により動作を指定するためのアドレスをアドレス/コマンド転送用バス(ACBUS)17上に送出する。

【0105】システムメモリ制御ユニット15が、I/O動作はプロセッサ1に対し向けられたものであり、リソースの争いは生じないことを確認した場合、このユニット15は、データチャンネル制御ユニット(DCB)16に対し、データチャンネルI/OD3からデータチャンネルI/OD1へデータの項目を転送するように命令する。これと同時に、データチャンネル制御ユニット16は、メモリから読み出されたデータを、メモリデータ入力/出力チャンネル(MDAT)19からデータチャンネルI/OD2へ転送するように命令される。

【0106】これまで述べてきた説明は、図11に示したように、データチャンネル制御ユニット(データ・クロスバー(DCB))16がデータ保持用素子を全くもっていない場合のアーキテクチャに言及している。それゆえに、データの項目の転送は、データチャンネル制御ユニット16を介し、一つの時間周期でもって行われる。

【0107】したがって、任意のプロセッサから出力されるべきデータは、システムバスに対しアドレスが提示される時間周期の後に続く時間周期において転送されるように決められている。この時間周期の差により、システムメモリ制御ユニット15は、要求されている動作に対応するリソースが使用可能であるか否かを検査するための時間が与えられる。

【0108】一つの時間周期において、データチャンネル制御ユニット16を介してデータが転送される際には、すべての長さのデータチャンネルI/OD(i)、データ

チャンネル制御ユニット16、および、メモリデータ入力/出力チャンネル19を通して上記の一つの時間周期内にデータが伝達されることが前提条件となる。この場合、クロック周期は、この前提条件により下限が決まる。

【0109】本発明の他の態様によれば、クロスバー相互接続ロジックは、入力保持レジスタと、出力保持レジスタとを備える。前者の入力保持レジスタは、受信部21、22、23、24および25のすぐ下流側の最初の位置に配置される。また一方で、後者の入力保持レジスタは、出力側のドライバ25、26、27、28および29の上流側の最初の位置に配置される。

【0110】入力保持レジスタのみを採用した場合、データバス、すなわち、データの経路は2つの流れに分割される。これらの流れの各々は、2つの連続するクロック周期に従って移動することが可能である。すなわち、この場合は、データが移動する全時間期間が、図3を参照したときに考えられる時間期間に等しくなるにしても、絶対的な時間は顕著に短くなる。

【0111】さらに、入力保持レジスタおよび出力保持レジスタを採用した場合、データバスはのみを採用した場合、データの経路は3つの流れに分割される。これらの流れの各々は、3つの連続するクロック周期の一つに従って移動する。上記のいずれの場合においても、メモリデータ入力/出力チャンネル(MDAT)19上で非常に高速の転送速度が得られる。さらに、異なるチャンネルからのデータ転送の位相に関し部分的な重ね合わせを行うことも実現される。

【0112】さらに、このようなデータバスの細分により、メモリとデータチャンネル制御ユニット(DCB)との間、および、データチャンネル制御ユニット(DCB)とプロセッサとの間で、異なるデータ転送の並列性を許容することができるようになる。この結果、個々のプロセッサの端子数を顕著に節減することが可能になる。図5は、図1のマルチプロセッサ・システムのデータ・クロスバーの好ましい具体的構成例を示す概略的なブロック図である。

【0113】ここでは、本発明の実施例のマルチプロセッサ・システムばかりでなく、本発明に関係する他の概念にも適用することが可能なデータ・クロスバー、すなわち、データチャンネル制御ユニットの構成をブロック形式で図示することとする。図5において、図2のブロック図で既に示した構成要素に対応する機能部分は、同一の参照番号を付して表すこととする。

【0114】図5に示すように、メモリデータ入力/出力チャンネル19は、64+8ビットにより構成される。このような構成においては、8バイトの並列形式(すなわち、二重のワード)で行われるメモリへのデータ転送、またはメモリからのデータ転送と、その後に続く8バイトのエラー訂正コード(ECC)とが可能になる。データチャンネル19は、受信部35およびドライバ29



に接続される。

【0115】受信部35の出力は、メモリから受信されるデータを保持するためのデータ保持用のレジスタ37に接続される。レジスタ37の出力は、シンドローム生成用ロジック(SYNDR GENと略記されることもある)38に接続されると共に、一般のタイプのエラー訂正ネットワーク(DATA CORRECTIONと略記されることもある)39に接続される。

【0116】シンドローム生成用ロジック38は、受信した情報を分析し、起こり得る訂正可能なエラーと、同様に起こり得る訂正不可能なエラーとを認識する(後者の訂正不可能なエラーに対しては、「エラー訂正不可能」を表示する出力信号を提示する)。さらに、上記のシンドローム生成用ロジック38は、訂正可能なエラーを訂正するために、エラー訂正ネットワーク39のロジックを制御する。

【0117】さらにまた、上記のシンドローム生成用ロジック38は、バイトの情報の各々に対し、パリティ制御用バイトに関連づける。このパリティ制御用バイトは、ロジックシンドローム生成用ロジック38により、エラー訂正ネットワーク39のロジックに転送される。エラー訂正ネットワーク39のロジックは、出力チャンネル40上に8バイトの情報を提供する。さらに、8バイトの情報の各々の後に、一つのパリティバイトが続く。

【0118】出力チャンネル40は、4つのグループのロジック回路41、142、143および144に対し情報を分配する。これらのロジック回路の各々は、プロセッサのデータチャンネルに連結される。これらの4つのグループのロジック回路41、142、143および144は、互いに同じタイプのものなので、ここでは、データチャンネルI/OD1に連結される1つのグループのロジック回路41のみを詳細に説明することとする。

【0119】ロジック回路41は、第1の72バイトのレジスタ42を備える。このレジスタ42の入力は、出力チャンネル40に接続される。また一方で、レジスタ42の出力は、18の要素からなる複数のグループとして、マルチプレクサ31に接続される。このマルチプレクサ31との接続により、レジスタ42の出力は、各々が18の要素からなる11のグループの入力を有するマルチプレクサ31の入力グループを形成する。

【0120】マルチプレクサ31の出力は、18のセルレジスタであるDO1用レジスタ44に接続される。このDO1用レジスタ44の出力は、ドライバ25の入力に接続される。このドライバ25の出力は、データチャンネルI/OD1に通ずる。マルチプレクサ31の4つのグループの入力45は、出力チャンネル40に直接接続される。

【0121】残りの4つのグループの入力は、それぞれ、各々が18本のワイヤからなる3つのチャンネル46、47および48に接続される。これらの18本のワ

イヤに対し、3つのグループのロジック回路142、143および144が、2バイトの情報と、これに続く一つのパリティバイトを供給する。上記マルチプレクサ31は、デコーダ36により生成される適切な選択信号によって制御される。このデコーダ36の制御機能によって、DO1用レジスタ44は、連続的なローディングが可能になる。そして、それゆえに、チャンネル40上に存在するか、または、レジスタ43内に保持されている二重のワードから抽出される一対のバイトの情報を、データチャンネルI/OD1へ連続的に転送することが可能になる。上記マルチプレクサ31はまた、データチャンネルI/OD2、I/OD3およびI/OD4から、それぞれロジック回路142、143および144を介して個々にやって来る一対のバイトの情報(および、関係するパリティバイト)を、データチャンネルI/OD1へ転送することも可能にする。

【0122】チャンネル40から二重のバイトを直接選択するマルチプレクサ31により提供される可能性によって、チャンネル40上に存在する二重のバイトに関しDO1用レジスタ44をローディングし、これと同時に、チャンネル40上に存在する二重のワードに関しレジスタ42をローディングすることが可能になる。このようにして、読み出し動作により明確にアドレス指定がなされた二重のバイトのプロセッサへの転送を、かなりの高速で行うことができる。

【0123】レジスタ42内に保持されている他の二重のバイトは、適切な順番で前者の二重のバイトの後に付加することができる。しかしながら、読み出し動作の流れは、さらに考察する必要がある。以下に、この読み出し動作の流れについてさらに詳しく述べることとする。メモリ内へのデータ書き込み、または、プロセッサ間のデータの転送のために、データ・クロスバーは、ロジック回路41のユニット内に、一つのグループの受信部21を備える。この受信部21の入力は、データチャンネルI/OD1に接続されており、その出力は、18のセルレジスタであるDI1用レジスタ(第1のレジスタ)49に接続されている。

【0124】DI1用レジスタ49の出力は、チャンネル50に接続される。このチャンネル50は、3つのグループのレジスタ(ロジック)142、143および144に対し、DI1用レジスタ49内に保持されている情報を分配する(特に、マルチプレクサ31と等価なマルチプレクサに対し)。DI1用レジスタ49の出力はまた、第2のレジスタ51に接続される。この第2のレジスタ51の出力は、第3のレジスタ52の入力およびパリティエラー検査用ロジック(PCHECKと略記されることもある)。

【0125】さらに、この第3のレジスタ52の出力は、第4のレジスタ54の入力に接続される。この第4のレジスタ54の出力は、キャッシュメモリ内の第5の

レジスタ55の入力に接続される。D I 1用レジスタ49およびレジスタ51は、18のセルをもっているのに対し、レジスタ52、54および55は、16のセルしかもっていない。この理由として、後者のセルにおいては、パリティビットを保持することは不必要であることが考えられる。

【0126】レジスタ51、52、54および55のバイト出力は、各々が64ビットからなる4つのチャンネルを有するマルチプレクサ56の第1のグループの入力57に接続される。他のグループの入力58、59および60は、ロジック回路41のグループに相当するロジック回路142、143および144にそれぞれ接続される。これらのロジック回路142、143および144は、それぞれ、データチャンネルI/OD2、I/OD3およびI/OD4に連結される。

【0127】マルチプレクサ56の出力は、8ビットコードのECC（既述のとおり、エラー訂正コードの略）生成用ロジック（ECC GENと略記されることもある）61の入力に接続される。このECC生成用ロジック61は、エラーを検出して訂正するためのものである。上記マルチプレクサ56の出力はまた、72ビットのレジスタ62の入力に接続される。このレジスタ62は、ECC生成用ロジック61により生成される8ビット入力形のECCのコードも受け取る。

【0128】レジスタ62の出力は、出力側のドライバ29の入力に接続される。このドライバ29は、メモリデータ入力/出力チャンネル（MDAT）19に通ずる。図6は、図5のデータ・クロスバーの動作を説明するためのタイミング図である。図6において、図4のタイミング図で既に示した信号の名前に対応する信号のラインは、図4の信号のラインと同じ意味を有する。

【0129】図6では、説明を簡単にするために、状態信号SNOOP OUTのダイアグラムを省略する。また一方で、DIREGのダイアグラムを、プロセッサに接続されたチャンネルの状態を表すデータチャンネルI/OD(i)のダイアグラムに付加している。このDIREGのダイアグラムは、データ・クロスバー（DCB）の入力側のレジスタ、例えば、レジスタ49およびレジスタ37の状態を表すものである。さらに、メモリデータ入力/出力チャンネル19の状態を表すメモリデータ入力/出力チャンネル（MDAT）のダイアグラムと、メモリからDCBへとデータが転送されるときの入力側のレジスタ37の状態を表すDOREGのダイアグラムと、DCBの出力側のレジスタ44の状態を表す状態DO(i)のダイアグラムとを、データチャンネルI/OD(i)のダイアグラムに付加している。

【0130】ここでは、データバスを複数の流れに分割することにより、例えば10nsecの非常に短いクロック周期（周期的信号CKの時間長に相当する）を用いることが可能になる。そして、ほんの2周期の間（各々

の転送に対し20nsec)だけアドレス/コマンド転送用バス（ACBUS）またはメモリデータ入力/出力チャンネル（MDAT）を占有することが可能になる。

【0131】さらに、各々の転送に対し、8バイトのデータ（2ワード）または多量のデータを転送することが可能になる。プロセッサのデータチャンネルのレベルでは、データの転送が、クロック周期の時間間隔で実行される時間毎に、部分的に連続する2バイトの転送形式で行われる。このようなデータの転送は、前述のアーキテクチャにおいて、各チャンネルが、使用可能のように調整されかつ「バッファ機能」を有するリソースをもっているという事実を利用することにより実行される。

【0132】このことは、幾つかの異なるデータチャンネルI/OD(i)とメモリとの間のデータ転送を時間的に重ね合わせる可能性をもたらす。このような可能性に加えて、システムのメモリデータ入力/出力チャンネル（MDAT）およびアドレスバスは、2種のノードを構成する。これらのノードにより、データおよびアドレスの連続的なかつ順序立った流れを重ね合わせることができ。さらに、データおよびアドレスに対し関連する相関関係のラベルを必要とすることなく、種々の動作の管理および制御が可能になる。この場合、このようなラベルは、余分なものとなる。

【0133】ついで、図6を順次考察していくこととする。図6において、一般的なプロセッサ1は、周期P1において、アクセス要求信号ABREQを提示し、周期P3において、システムバスおよびデータチャンネルに対するアクセスの許可を受け取る。さらに、周期P5および周期P6において、プロセッサ1は、アドレスとこのアドレスに関係するコマンドを、アドレス/コマンド転送用バス（ACBUS）上に出送する。

【0134】さらに、周期P8および周期P9において、システムメモリ制御ユニット15からメモリアドレス用チャンネル（MADDR）へアドレスが転送される。この間に、プロセッサ1は、周期P5において、二重のバイトのデータをデータチャンネルI/OD1上に出送する。このようにして送られたデータは、周期P6において、D I 1用レジスタ49（図5）内で保持される。そして、このようにして保持されたデータは、その後の周期において、D I 1用レジスタ49からレジスタ51、52、54および55へ漸次転送される。

【0135】周期P10において、データチャンネルI/OD1を介して受信された最初の二重のバイトのデータは、レジスタ55内に保持される。周期P10において、プロセッサ1は、2番目の二重のバイトのデータをデータチャンネルI/OD1上に出送する。この2番目のデータが、D I 1用レジスタ49からカスケード接続のレジスタ51、52、54へ転送されると、この転送されたデータは、周期P10の期間よりレジスタ54内に保持される。

【0136】同じようにして、プロセッサ1は、周期P7および周期P8において、3番目4番目の一対のバイトのデータをデータチャンネルI/OD1上に送出する。このようにして送出されたデータは、それぞれ、レジスタ51、52内に保持され、周期P10の期間より使用可能になる。このようにして、プロセッサ1は、4つの周期P5～P8の期間で、連続する8バイトの対のデータの転送を実行する。さらに、周期P10の期間より、マルチプレクサ56の出力において、8バイトのデータが並列形式で使用可能である。

【0137】周期P12およびP13において、マルチプレクサ56がイネーブルの状態になり、情報がレジスタ62に転送される。このレジスタ62によって、転送された情報が保持され、かつ、メモリデータ入力/出力チャンネル(MDAT)上に出力データが保持される。周期P3において、他のプロセッサ2がアクセス要求信号ABREQ2を提示し、周期P5において、プロセッサ1により既に使用されているモジュールとは異なるモジュールの書き込み動作に関するバス許可信号ABGRANT2を受け取る。そして、必要なリソースが空いている場合、プロセッサ2は、周期P7および周期P8の期間で、アドレス/コマンド転送用バス(ACBUS)上にアドレスを送出し、かつ、周期P7～周期P10の期間で、データチャンネルI/OD(2)上に4対のバイトのデータを連続して送出することにより、書き込み動作を開始して完了させることができる。

【0138】このような情報は、周期P14および周期P15において、コピーされた後にレジスタ62内に保持される。それゆえに、2つのプロセッサ1、2からメモリへの転送は、部分的な時間の重ね合わせにより行われる。読み出し動作は、書き込み動作とほぼ同じ流れでもって進行する。

【0139】例えば、周期P5の期間でプロセッサ1から提示されたアクセス要求により、周期P7において、アクセスの許可が得られる。このアクセスの許可により、周期P9および周期P10の期間において、アドレス/コマンド転送用バス(ACBUS)がアドレスによって占有される。さらに、周期P12および周期P13において、アドレスは、メモリアドレス用チャンネル(MADDR)18に転送される。

【0140】例えば、周期P20および周期P21において、読み出されたデータの項目は、メモリデータ入力/出力チャンネル(MDAT)19上で使用可能であり、周期P21および周期P22において、レジスタ37(DOREGのダイヤグラム)内に保持される。周期P22において、マルチプレクサ31およびレジスタ42は、一対のバイトのデータをDO1用レジスタ44に転送し、かつ、メモリから受信される8バイトのデータをすべてレジスタ42内にローディングするように制御される。

【0141】さらに、周期P22において、メモリデータ入力/出力チャンネル(MDAT)19およびレジスタ37が空いた状態になるので、これらのチャンネル19およびレジスタ37は、例えば他のプロセッサにより予め定められた別の情報を転送して保持することができる。周期P23において、DO1用レジスタ44内に保持された二重のバイトのデータは、データチャンネルI/OD1上に転送され得る。また一方で、上記のDO1用レジスタ44では、レジスタ42内に保持されたデータの中から、マルチプレクサ31により選択された二重のバイトのデータがローディングされる。

【0142】周期P24、周期P25および周期P26において、その後続く3対のバイトのデータが、データチャンネルI/OD1上に転送され、転送動作が完了する。この場合、上記の読み出し動作が、他の読み出し動作と部分的な重ね合わせがなされるような転送動作によって遂行することができることは、明らかなことである。

【0143】例えば、周期P3において、書き込み動作よりもむしろ読み出し動作に関係するアクセス要求が、プロセッサ2により提示された場合、周期P7および周期P8の期間でリソースが使用可能であるという仮定が成り立つときは、周期P18および周期P19において、読み出すべきデータの項目がメモリデータ入力/出力チャンネル(MDAT)上に存在し、さらに、周期P19および周期P20において、上記データの項目がレジスタ37(DOREGのダイヤグラム)内にローディングされるであろう。

【0144】周期P20～周期P23の期間で、DO1用レジスタ44(DO(1)のダイヤグラム)へのブロック転送が行われる。さらに、周期P21～周期P24の期間で、データチャンネルI/OD2への転送が行われる。このデータチャンネルI/OD2への転送は、DO1用レジスタとデータチャンネルI/OD1の進行の動作に関する部分的な時間の重ね合わせによって実行される。

【0145】ここで、周期P9において、プロセッサ3が、書き込み動作のためのシステムバスへのアクセス要求を行うと仮定する。この場合、上記のアクセス要求が提示された周期から10クロック周期後にメモリデータ入力/出力チャンネル(MDAT)19が使用可能であることが前提条件となる。すなわち、メモリデータ入力/出力チャンネル(MDAT)が、一方で、周期P5の期間で差し出されたアクセス要求を満足するように定められているような、周期P20および周期P21において、MDATが使用可能でなければならない。

【0146】それゆえに、一度、システムメモリ制御ユニット15がバスへのアクセスを許可し、対象とする動作を書き込み動作として認識すると(周期P13および周期P14)、このシステムメモリ制御ユニット15は、メモリアドレス用チャンネル(MADDR)18上へ

のアドレスの転送を阻止することにより、トランザクションを中断させる（周期P16）。さらに、上記のユニット15は、メモリデータ入力/出力チャネル（MDAT）19上へのデータの転送を阻止する。その後、上記のユニット15は、出力信号ARESPとして予め定められた期間（周期P18および周期P19）に提示されるリトライ信号RETRYを用いて、プロセッサ3に対し、周期P21またはその後の周期で強制的にアクセス要求を繰り返させる。

【0147】それゆえに、一方では、データ転送動作が、メモリ内に書き込むために9クロック周期を必要とし、かつ、メモリから読み出すために18クロック周期を必要とし、他方では、2種の転送動作間の干渉時間、および、2種の転送動作間で起こり得る衝突時間が、ほんの2クロック周期に限定されることが認められる。このために、部分的に時間の重ね合わせがなされた転送が可能である。このような転送は、各種のメモリ・リソース（モジュール）と、各種のプロセッサのチャネル（データチャネルI/OD(i)）とを使用することにより実行される。これに加えて、上記の転送は、これらのチャネルに係るバッファとしてのリソース、直列式のリソース、および並列式のリソースを使用することによっても実行される。これらのリソースは、データ・クロスパー（DCB）のロジック内のプロセッサのチャネルに連結される。

【0148】さらに、図5のブロック図および図6のタイミング図に基づき、モディファイされたデータの項目を他のプロセッサへ転送する際にあるプロセッサによる介在があった場合でも、転送動作は、対のバイトのデータをシリアル形式にして直接に行われることが、すぐさま結論づけられる。さらに詳しくいえば、この転送動作では、機能的に図4のD11用レジスタ49と等価なレジスタから、チャネル50、46、47および48の一つを介して、機能的に図4のD01用レジスタ44と等価なレジスタへと対のバイトのデータが転送される。

【0149】図3のタイミング図を参照しながら既に説明したように、上記の転送動作は、全体として、プロセッサおよびメモリ間の一つまたは2つ以上の転送に対し時間的に重ね合わせられる。今までは、本発明の特定の好ましい実施例についてのみ説明してきたが、多くの適切な変形例が考えられることは明らかである。

【0150】プロセッサおよびメモリモジュールの数（好ましい実施例においては、4つのプロセッサと、6つのモジュール）は、メモリの並列性とプロセッサの並列性との間の並列性の比が任意に設定されるように選択することができる。多数の並列性を達成するために、さらに多くのデータ・クロスパー（DCB）のロジック構成要素が並列形式で使用され得る。また、この場合、データ・クロスパーのロジックは、パリティ検査回路以外に、エラーの検出および訂正のためのエラー訂正部およ

びコード生成部を備えている。さらに、データ・クロスパーのロジックは、メモリから読み出される情報と、メモリ情報の部分的なモディファイのためにプロセッサからやって来る他の情報とを結合する（マージングともいう）ための回路も備えている。

【0151】さらに、アドレスバスおよびコマンドバスに対するアクセス（ABREQ(i)）と、データチャネルに対するアクセス（DBREQ(i)）のアービトレーションのために各々独立の信号を使用することも可能である。このアービトレーションでは、読み出し動作/書き込み動作またはその他の動作に関するアクセスの要求を提示するためのトランザクションや、存在するかまたは予定されている必要なリソースの使用可能性に対し、バスの許可を調整するためのトランザクションを特徴としている。このような構成により、「リトライ」のケースを最小限に減らし、それゆえに、システムバスの最適な利用を実現することができるようになる。

【0152】同じプロセッサが、連続的な読み出し要求の後にかかなりの程度の連続性でもってデータを受け取ることを可能にするために、レジスタ42は、複数のレジスタをカスケード接続形式にするか、または、FIFO（First in-First out）のスタック形式にすることによって構成することもできる。リソースの争いがあるような複数種の書き込み動作の場合は、リトライ動作を回避するために、同じような概念が使用される。さらに詳しくいえば、このような概念は、図5のレジスタ51、52、54および55の下流側に配置される入力バッファに各種の書き込み動作を保持すると共に、アドレスを一時的に記憶するための同じような入力バッファをシステムメモリ制御ユニット15内に設けることにより使用される。

【0153】このようにすれば、予め定められた周期の期間内に遂行できない書き込み動作は、必要なリソースが使用可能になるようなその後の周期にまで延長することができる。さらに、複数のプロセッサのすべてまたはその一部が、キャッシュメモリを備えることは必ずしも不可欠なことではない。この理由として、本発明の主題を構成するアーキテクチャにより提供される利点は、プロセッサ間のデータ転送が、プロセッサおよびメモリ間の転送の重ね合わせにより実行されるという事実によって達成されることが挙げられる。

【0154】最終的に、これまでの説明において、一つのグループのプロセッサ、すなわち、「一群のプロセッサ」をも包含させるように、「プロセッサ」という用語を使用することができることを明らかにしなければならない。これらのプロセッサは、局所的バスと相互接続がなされ、かつ、システムバスと通信すると共に、インタフェース用アダプタを介してのデータの転送を行うための二地点間チャネルと通信することができる。このような構成により、外部的な効果を考慮した場合、一群のプ

ロセッサが、単一のプロセッサとみなせるようになる。

【0155】この場合、幾つかのグループのプロセッサをシステムバスに対し直接に接続するというインタフェース用アダプタなしの直接接続も可能である。グループ内の各プロセッサは、同じデータ転送チャンネルに対し直接に接続される。このデータ転送チャンネルは、幾つかのプロセッサとの接続に関しては分岐データバスと考えられ、また一方で、プロセッサの集合体に関し、ならびに、データチャンネル制御ユニット（データ・クロスバ）16との接続に関しては二地点間データバスと考えられる。

【0156】この場合、当然のことではあるが、データの「転送速度」は、より低くなるであろう。この理由として、データチャンネルに対する負荷が比較的大きくなることが挙げられる。そして、周期的信号CKの周波数を比較的低い値に設定することが必要になるであろう。二者択一的な例として、第1の複数のプロセッサの各々が、複数の二地点間データチャンネルを介してデータ・クロスバと通信し、また一方で、第2の複数のプロセッサ（比較的低速で済む周辺制御装置として機能するプロセッサ）の各々が、単一の分岐（データチャンネル）バスを介してデータ・クロスバと通信するようなシステムの場合を考える。この場合、このバス上のデータ転送は、転送周波数を変えないままにした状態で（転送される各ブロックに対し一つのクロック周期で）幾つかのクロック周期にわたり（例えば、2クロック）バスを占有することによって実行される。

【0157】上記の解は、明らかに、データ・クロスバが、図5に示したようなタイプ、すなわち、バッファ用レジスタを備えたタイプである場合にのみ好都合である。図7は、図1のマルチプロセッサ・システムを变形した実施例を示す概略的なブロック図である。図7においては、図1と同じように、本発明のマルチプロセッサ・システムのアーキテクチャが概略的に図示されている。ここでは、先の図1に示した構成要素と機能的に等価な構成要素に関しては、同一の番号を付して表すこととする。

【0158】図7のブロック図は、プロセッサ1、2が、一対のプロセッサにより構成されているという事実のみにおいて図1のブロック図と異なる。図7では、プロセッサ1は、2つのプロセッサ101、102により構成される。これらの2つのプロセッサ101、102は、アドレス／コマンド転送用バス（ACBUS）17およびデータチャンネルI／OD1に対し直接に接続される。

【0159】さらに、これらの2つのプロセッサ101、102は、アービトレーション用のシステムメモリ制御ユニット15から見れば、互いに競合する2つの独立したプロセッサとみなせる。この場合、2つのプロセッサ101、102は、コマンドおよびアドレスバスに

に対するアクセスばかりでなく、データチャンネルI／OD1に対するアクセスをも考慮したときに、2つの独立したプロセッサであると考えられる。

【0160】システムメモリ制御ユニット15は、アービトレーション用ユニットおよび有限状態ロジックの両方に対し上記の事実を考慮に入れる。ここで、上記の2つのプロセッサ101、102が、時間的な重ね合わせによりデータチャンネルI／OD1上でトランザクションを実行できないことは明らかである。プロセッサ2は、2つのプロセッサ103、104と、インタフェース用アダプタであるインタフェース・ロジック105とにより構成される。

【0161】これらの2つのプロセッサ103、104は、互いに通信し合い、かつ、一般的なタイプの局所的バス106を介してインタフェース・ロジック105と通信する。このインタフェース・ロジック105は、アドレス／コマンド転送用バス17およびデータチャンネルI／OD2に接続される。さらに、インタフェース・ロジック105は、局所的バス106に対するアクセスのアービトレーション、すなわち、調整を行う。このアービトレーションは、システムバス（アドレス／コマンド転送用バス）17およびデータチャンネルI／OD2に対し2つのプロセッサ103、104により提示されるアクセス要求を認識することにより実行される。

【0162】これらのアクセス要求は、システムバスのプロトコルおよびタイミングに従って、このシステムバスに転送される。局所的バス106が非同期タイプであり、プロセッサ103、104の動作が非同期形式で遂行されることは明らかである。これに対し、インタフェース・ロジック105が、システム内の他の構成要素と同期して動作するように、周期的信号CKにより時間規定がなされなければならないことは明らかである。上記のプロセッサ103、104は、システムバス（アドレス／コマンド転送用バス）17およびデータチャンネルI／OD2に対し直接に通信するので、同じ条件下にあることが必要である。

【0163】この場合、プロセッサ103、104は、システムメモリ制御ユニット15により単一のプロセッサとみなされる。そして、インタフェース・ロジック105は、一方のプロセッサ、または他方のプロセッサに対し、受信したメッセージデータを振り分ける役割を遂行する。今まで本発明の特定の実施例について説明してきたが、ここでは、ただ単に、本発明のほんの一例を例証したにすぎないと考えられる。さらに、当業者においては数多くの変形および変更が容易になし得るので、本文で示したような構成にのみ本発明を限定することは望ましくない。したがって、本文に添付されている請求の範囲およびその等価物に記載された発明の範囲内にある限りにおいては、すべての適切な変形例および等価例が考えられる。

## 【図面の簡単な説明】

【図1】本発明の一実施例に従って構成されるアーキテクチャおよび共用メモリを有するマルチプロセッサ・システムを示す概略的なブロック図である。

【図2】図1のアーキテクチャのデータチャンネル制御ユニットの具体的構成例を示す概略的なブロック図である。

【図3】図1のアーキテクチャのシステムメモリ制御ユニットの具体的構成例を示す概略的なブロック図である。

【図4】図1のマルチプロセッサ・システムの動作を説明するためのタイミング図である。

【図5】図1のマルチプロセッサ・システムのデータ・クロスバの好ましい具体的構成例を示す概略的なブロック図である。

【図6】図5のデータ・クロスバの動作を説明するためのタイミング図である。

【図7】図1のマルチプロセッサ・システムを変形した実施例を示す概略的なブロック図である。

## 【符号の説明】

20

1, 2, 3および4…プロセッサ

5…システム・メモリ

10, 11, 12, 13, 113および114…モジュール

14…タイマ・ユニット

15…システムメモリ制御ユニット

16…データチャンネル制御ユニット

17…アドレス/コマンド転送用バス

18…メモリアドレス用チャンネル

19…メモリデータ入力/出力チャンネル

31…マルチプレクサ

37, 42, 51, 52, 54, 55および62…レジスタ

41, 142, 143および144…ロジック回路

44…DO1用レジスタ

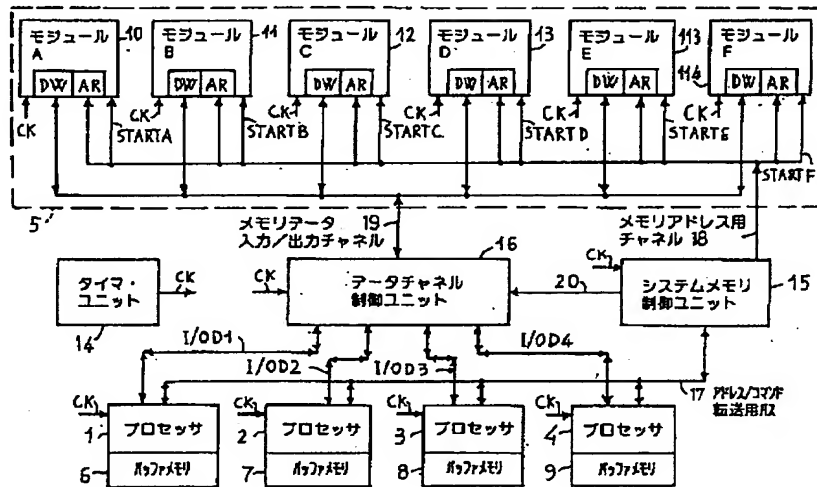
49…DI1用レジスタ

56…マルチプレクサ

70…アービトレーション用ロジック

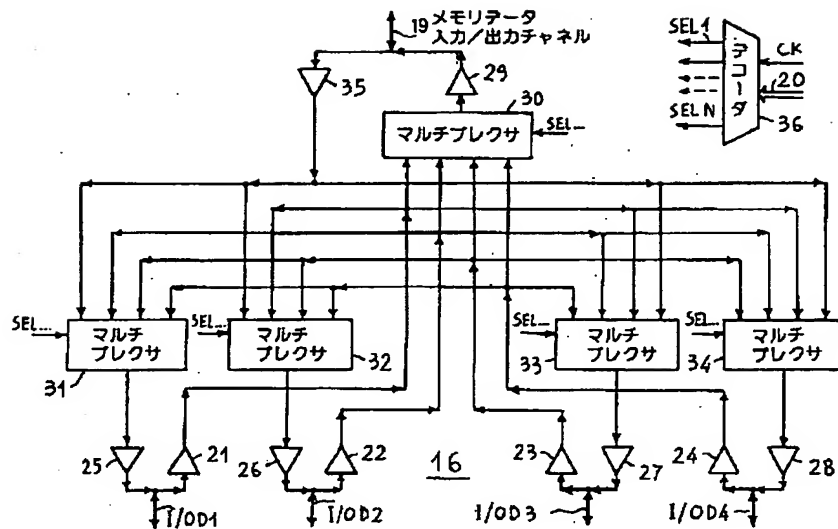
72…有限状態ロジック

【図1】

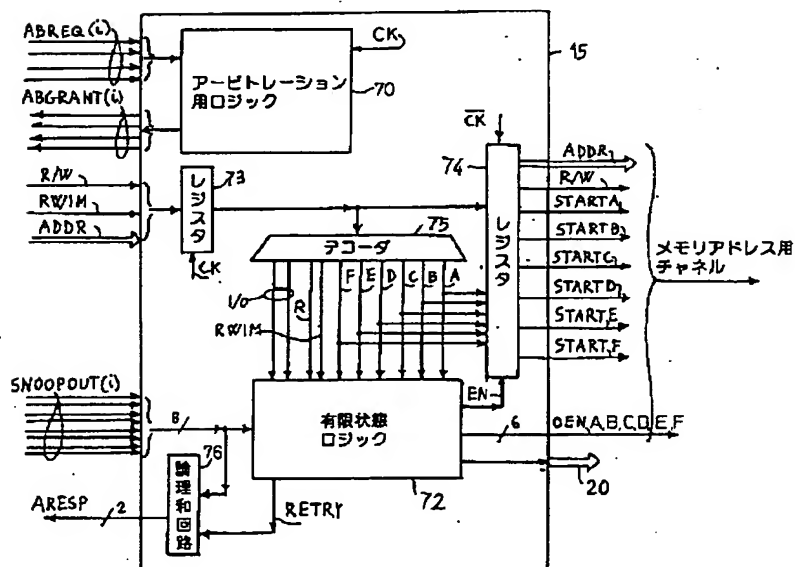




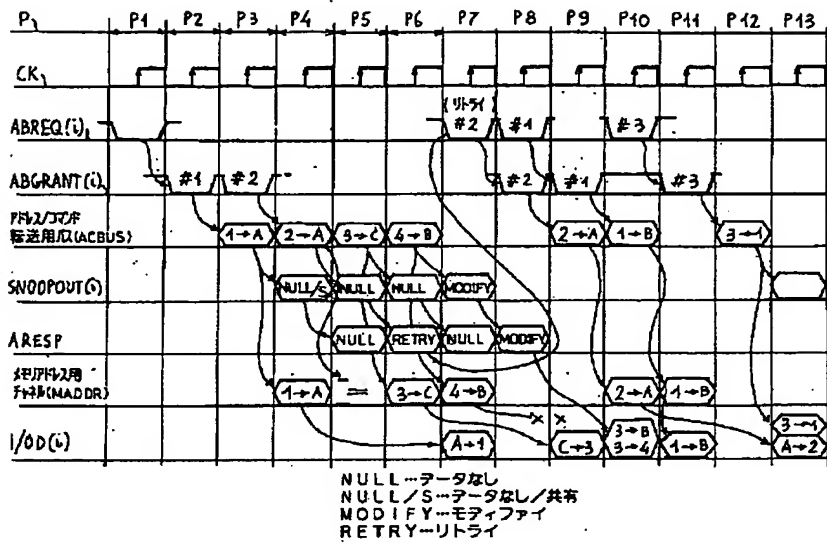
【図2】



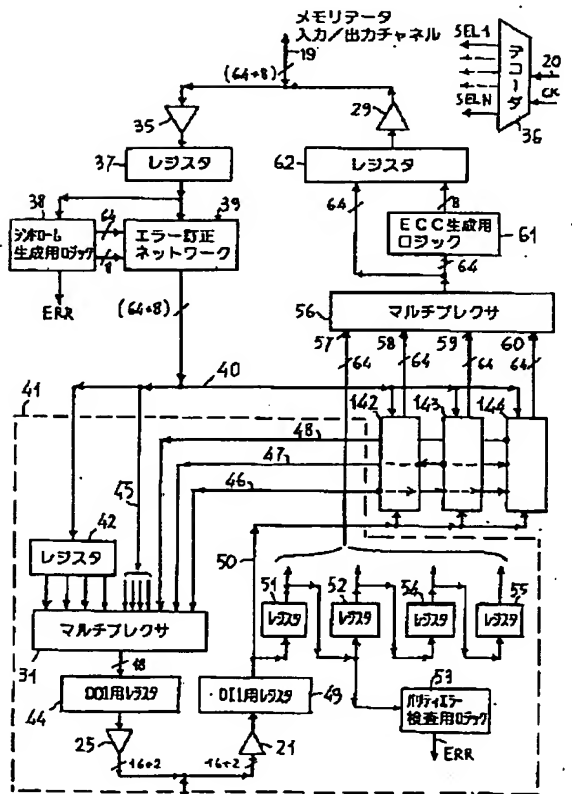
【図3】



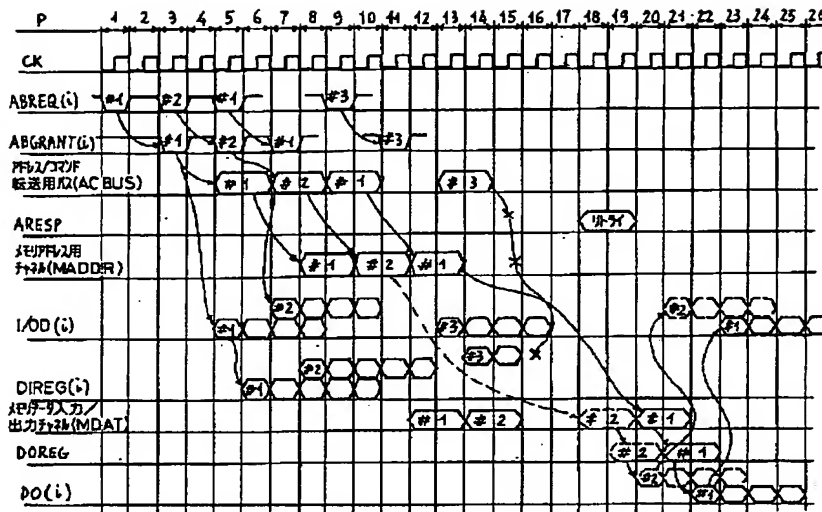
【図4】



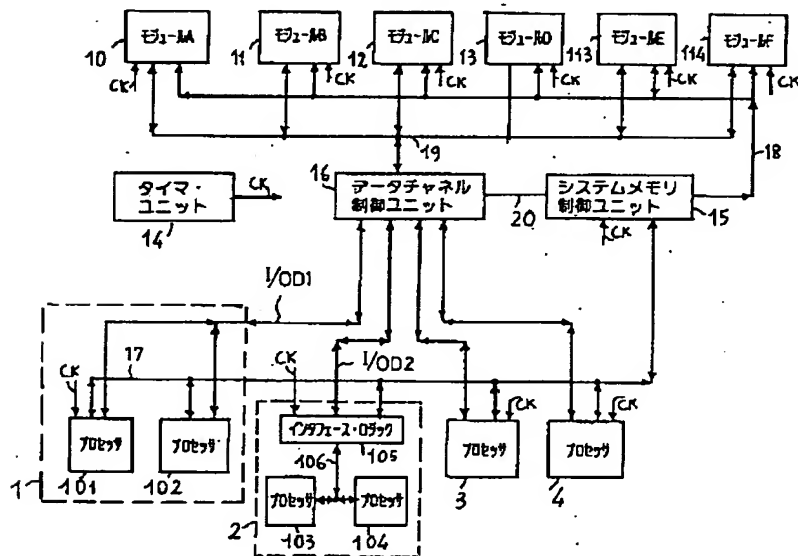
【図5】



【図6】



【図7】



フロントページの続き

(51) Int. Cl. 6

G 0 6 F 13/12

識別記号

庁内整理番号

F I

技術表示箇所

3 1 0 J

7368-5B

(72) 発明者 カルロ バニョーリ

イタリア国, 20131 ミラノ, ピア パラ  
ツツエ, 78

(72) 発明者 アンジェロ ラツァリ

イタリア国, バビア, 27029 ビジェバノ,  
ピアレ ベトラルカ, 20